Matrix Equation

Definition

A matrix equation has the form

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

where $\mathbf{A}$ is a matrix, $\mathbf{x}$ and $\mathbf{b}$ are column vectors.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

This can be explained as a system of linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Gauss-Jordan Reduction

When we handle a linear equation system, it is common practice to multiply an equation by a constant and to add or subtracts two equations to eliminate variables. In the end, we want to get something like

$$\begin{aligned} x_1 &= r_1 \\ x_2 &= r_2 \\ &\vdots \\ x_n &= r_n \end{aligned} \tag{1}$$

where $r_1 \dots r_n$ is our solutions. The whole process can be done in the language of matrix.

First of all, we need to write down the linear system in form of a matrix. It is obviously that we should not only record the information from the left-hand side but also that from the right-hand side. So, we have the augmented matrix

$$
[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]
$$

Then, we introduce elementary row operations to describe the common practices to elimininate variables:

- Interchange two rows.
- Multiply a row by a constant $c \neq 0$
- Add c times row j to row i, leaving row j unchanged.

And, we rewrite (1) as an augmented matrix,

$$
\left[\begin{array}{cccc|c} 1 & 0 & \dots & 0 & r_1 \\ 0 & 1 & \dots & 0 & r_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & r_m \end{array}\right]
$$

which called the row echelon form.

In summary, to solve the original algebraic system, we can write down the system as an augmented matrix and use elementray row operations to transform it into the row echelon form, from which we can read the solution(s) directly. That's the basic workflow of the Gauss-Jordan reduction.

Notes:

- In the row echelon form, each column with just 1 is called a pivot column.
- The row echelon form shown above it just the ideal case. In some case, the system may be reduntant and some of the rows turn out to be all-zero. Also, if the system is underdetermined, there can be some non-zero value following the pivots. These columns represent free variables.
- Although the elementary row operations provide a way to reduce all linear systems, directly using them can be complicated. Thus, we can add one enhanced row operation: given several rows $r_1 \dots r_n$, we can compute a new row $w = c_1 r_1 + \cdots + c_n r_n$ to replace one of the rows from $r_1 \dots r_n$, given that $c_1 \dots c_n$ are non-zero constants.